

Redefining Filter Bubbles as (Escapable) Socio-technical Recursion

Huw C. Davies

Oxford Internet Institute, University of Oxford

Key Words

Filter bubbles, technological determinism, digital sociology, algorithms, platform capitalism

To get the best of Google you have to know about what you don't know: what to look for and how to find it.

Abstract

Personalisation of media content is not a new phenomenon. Now, however, by configuring our search results and data feeds, algorithms that 'learn' from our digital footprint are determining what we see and hear. Pariser (2011) calls this the "Filter Bubble Effect". Yet, despite concerns that this effect is a threat to deliberative democracy we are told there is relatively little evidence to substantiate its existence (see Zuiderveen Borgesius et al., 2016). This article draws on a case study to argue that this is because the existing research looks for technical effects while neglecting our social lives. If we follow Foucault's reasoning that systems of thought are also technologies (Foucault, 1988) (Behrent, 2013) then we can see that material technologies (or what Foucault (ibid) called 'technologies of production') and immaterial technologies (ideas formed in discourse) can co-constitute filter bubbles. Borrowing language from computing and STS, this leads to a redefinition of filter bubbles as socio-technical recursion. This case study illustrates just one potential combination of such material and immaterial technologies (namely, search engines

and ideas that are encountered and formed during an individual's social life within their culture and class) that can create socio-technical recursion. The article concludes by arguing the advantage of conceptualising filter bubbles in this way is that it offers us a theoretical foundation for breaking out of this recursion by simultaneously challenging the mediums and messages that sustain them.

Introduction

The “filter bubble effect” (Pariser, 2011) is a metaphor for technology that only serves us information that is consistent with what we already know and feel about our world. Ostensibly, is not a new concept: just reading the same politically affiliated newspaper is the prototypical form of ‘bubbling’ (the practice of creating a filter bubble). In 1995, Negroponte predicted that the World Wide Web would amplify bubbling. He called it the “daily me” effect – that produces a “kind of newspaper printed in an edition of one” (Negroponte, 1996: 153). By 2006, Sunstein was saying the Web had allowed us to create our own information cocoons (Sunstein, 2006). Building on this, Pariser (2011) identified a significant development that has encouraged us to reconsider what we know about filter bubbles: namely the proliferation of machine learning technology that gives search engines, news aggregators, and social networks their ability to use records of our previous interactions with them to preselect and order what we see. It is claimed that this technology has co-opted, systemised, and intensified our habit of personalising our media content and consumption (Flaxman, Geol, & Rao, 2016). In response, anxiety is building. Tom Steinberg, a former advisor to the UK government and founder of MySociety.org, has recently stated that: “This echo-chamber problem is now SO severe and SO chronic that I can only beg any friends I have who actually work for Facebook and other major social media and technology to urgently tell their leaders

that to not act on this problem now is tantamount to actively supporting and funding the tearing apart of the fabric of our societies” (Steinberg, 2016). Similarly, the High Level Expert Group on Media Diversity and Pluralism, an independent group advising the European Commission, has stated that filter bubbles “will tend to create more insulated communities as isolated subsets within the overall public sphere. (...) Such developments undoubtedly have a potentially negative impact on democracy” (Zuiderveen Borgesius et al., 2016: 27). Recent research from Pew in America and Ofcom in the UK is compounding this anxiety. This research tells us that people are increasingly getting their news from social media platforms (Pew Research Centre, 2016); especially the young (Ofcom, 2013). And, data provided by Google’s News Lab tells us that when people want to research topics they have read about in their news feeds they will begin by consulting a search engine (Mitchell, Gottfried, Shearer, & Lu, 2017).

There are three main concerns about filter bubbles. Firstly, that search engines make an assessment of their user’s preferences and search history to customise their results for each individual (Pariser, 2011). Secondly, because social networks are homophilous (McPherson, Smith-Lovin, & Cook, 2001) (i.e. birds of a feather flock together), their users will only see news and opinion that is distributed by users in their network who share their political worldview. And thirdly, that social network platforms such as Facebook use algorithms to draw on data from these homophilous networks, and combine it with data from a user’s digital footprint to amplify the filter bubble effect by providing them with personalised newsfeeds. Yet, a review of the literature shows that there is little unequivocal evidence to justify these concerns (Zuiderveen Borgesius et al., 2016).

Flaxman, Goel and Rao (2016) looked for filter bubbles in news flowing through social media. They started by analysing the web-browsing behaviour of 1.2 million US-located users for the three-month period between March and May 2013, and eventually they focused on 50,000 users who actively read the news. They conclude: “news articles found via social media or web-search engines are indeed associated with higher ideological segregation than those an individual reads by directly visiting news sites” (Flaxman, Goel, & Rao 2016: 318). Conversely, however, they say that users of social media are also more likely to encounter counter-normative information because these information channels are “associated with greater exposure to opposing perspectives” (Flaxman, Goal & Rao 2016: 318). In their analysis of 10.1 million Facebook users, Bakshy, Messing, & Adamic (2015) also found that “individuals are exposed to more cross-cutting discourse in social media than they would be under the digital reality envisioned by some” (Bakshy et al., 2015: 1131). From this research, it is therefore difficult to conclude whether social media users see less or more partisan information as a consequence of being online.

An on-going research project in Germany is using a browser plugin to allow users to donate their search results for research. The plugin does not collect the actual searches of users, but automatically conducts searches for a fixed list of 16 terms, including, for example, “Angela Merkel”. So far, close to 4,000 contributors to the experiment have produced over 2 million queries. In quantifying the filter bubble effect the research team concludes that approximately 1-3% of users in the sample get “very different” results when searching for the query terms (Puschmann, 2017). Again, the results suggest that technological effects appear to be relatively weak.

Dutton, Reisdorf, Dubois, & Blank (2017) claim that the evidence for an effect has been lacking because the filter bubble hypothesis is a form of technological

determinism. They say that research often pays too much attention to what happens online at the expense of “on-going social process that is not limited to a single platform or type of interaction at one point in time” (ibid: 12). They therefore argue in our research methodologies we should acknowledge that “individuals seek and share information from and with a variety of sources and via a range of communication channels both online and offline (Chadwick, 2013), and over an extended period of time” (Dutton et al., 2017: 12).

Drawing on a case study, this article argues that we should follow this reasoning and address people as actors situated at the intersection of technology, culture and class. This doesn't simply exchange technological determinism for cultural or class determinism: it doesn't mean people's views are fixed or that they believe everything they read online, or that they only read what they want to believe. It does, however, mean that people's experiences and understanding of society and technology are reflected in their digital practices (for example in their use of search engines), and that this has consequences for the content of their information ecospheres. This move beyond technological determinism is achieved by broadening the definition of technologies to include ideas that emerge from our social lives.

Foucault (1988) claimed there are four forms of technology. Firstly, “technologies of production, which permit us to produce, transform, or manipulate things” such as search engines; secondly “technologies of sign systems, which permit us to use signs, meanings, symbols, or signification”; thirdly, “technologies of power, which determine the conduct of individuals and submit them to certain ends or domination, an objectivizing of the subject”; and finally “technologies of the self” (ibid: 17). Broadening the definition of technologies to include these other technologies introduces the social contexts including culture and class from which

these technologies emerge. Eventually, all this accounts for the many axes of difference that were in the data but were not attributable solely to “the technologies of production” (ibid: 17).

Study Design

This case study located two groups of young people aged 16 to 18 in two sharply contrasting institutions. One group was from a multi-ethnic, all-boys, fee-paying school in London, England that I have called The Chancery School and another group was from a publicly-funded, co-educational college of further education that I have called Calshot College, with a primarily white working class intake from a city on England’s south coast. The Chancery students were all in their first year of A levels (a post-16 qualification recognised by universities in the UK). The Calshot students were studying for a BTEC in Media Production, which is also a post-16 qualification but it includes vocational components (it is therefore less likely to be recognised by universities that demand high A level grades from their undergraduates).

The Chancery School’s culture is synchronised with London’s corporate elite. The signifiers of London’s mercantile power and heritage such as Latin mottos, heraldic coats of arms, and gilded portraits of patriarchs are combined at the school with polished chrome and glass behind which young men wearing business suits used tablets in class. The Chancery School students in the research cohort were from a range of ethnic and national backgrounds including English, Indian, Canadian, Jewish, Egyptian, and Burmese. They had all recently been offered a place at an elite university. At the Chancery School all the parents were professionals, many of them in prestigious roles, including a healthcare consultant, a university professor, and a high court judge (see Table 1).

Table 1:

The Chancery students' parental occupations and intended next steps

Name	Father's occupation	Mother's occupation	Intended Next Step
Omar	Professor	Admissions tutor	Oxford - Computer Science & Philosophy
Caleb	GP	Practice manager	UCL - Medicine
Zac	Lawyer	Lecturer	UCL - English
Daniel	Architect	Unavailable	Princeton
Tariq	Doctor	Computer scientist	Kings College / Imperial - Medicine
Ryan	Civil Judge	Professor	Cambridge - Human Social & Political Science
Tom	Retired consultant	Haematologist	Oxford - Classics
Stephen	University senior management	Professor	Oxford - Classics
Seb	Financial professional	School nurse	Imperial - Biology Chemistry & Maths
Saul	Endocrinologist	GP	Gap year working in research lab
William	IT consultant in banking	Linguist	Imperial - biological engineering

In contrast, Calshot College is situated in a relatively deprived area. Its modern office block architecture stands out among decaying nineteenth century municipal buildings and warehouses once used for industry. Its semi-circular sentinel reception booth, soft 'break-out' areas, and large boardrooms with plasma screens, high backed chairs and blinds also suggests that the college is also attempting to acclimatise its students to a particular corporate culture so that the college and students are synchronised in their aspirations. At Calshot College the cohort was a mixture of young men and young women. They all identified as white British or English, and they all lived locally. With one exception, all the students' parents were

unemployed, or employed in technical, engineering, semi-skilled or unskilled service occupations. (see Table 2).

Table 2:

The Calshot students' parental occupations and intended next steps

Name	Father's occupation	Mother's occupation	Intended Next Step
Liam	Air steward	Sales representative	University or job
Alfie	Sales administrator	Post office clerk	Job
Sarah	Call centre operator	Housewife	Staying at Calshot
William	Unavailable*	Unavailable	Job
Jake	Unavailable	Unavailable	Job
Phil	Mechanical engineer	Cleaner	University or job
Jessica	Unavailable	Cleaner	Undecided, probably job
Paul	Unavailable	Unavailable	Unavailable
Mel	Mechanical engineer	Cleaner	Undecided
Alex	Unavailable	Unavailable	Unavailable
Lizzy	Naval communications engineer	Charity worker	University
Beth	Unavailable	Teacher	Job
Rob	Health and Safety in building trade	Retail	Stay on or career in music
Josh	Unavailable	Unavailable	Unavailable
Matt	Shop assistant	Hospital assistant	Job

(*Unavailable means the student either didn't know what their parents did or they felt uncomfortable disclosing this information)

We know that parental education can be a useful proxy for class (Erola, Jalonon, & Lehti, 2016). Generally, the Calshot students were either working or lower middle class while the Chancery students were either upper middle or upper class. The main

objective here was to operationalise a combination of culture and class in order to compare the way the two cohorts engaged with information online via search engines.

Besides face-to-face communication, many young people now use an array of devices to communicate (including in game chat on consoles, smartphones and tablets (Ofcom, 2016)) and they use a variety of platforms to inform themselves about the world, including TV and social media (Ofcom, 2016). We should therefore observe users across platforms when examining their communication environments (Dutton et al., 2017). However, I have chosen here to focus on young people's use of search engines for two reasons. Firstly, following young people, particularly across social media platforms, is ethically problematic. The researcher would need their permission to access their Facebook feeds and Twitter networks, for example. Platforms popular with teenagers such as WhatsApp are encrypted. Since these data sets might potentially contain sensitive information about non-consenting participants under eighteens, achieving ethical clearance for this would be, at the very least, challenging. Secondly, for young people, Google is an increasingly important gateway to knowledge. According to research by the UK's media regulator Ofcom (2016): 'While the BBC remains the preferred source of 'true and accurate information about things that are going on in the world' among young people (at 35%), this is much less likely than in 2015 (52%). Young people are now more likely to choose Google in their response to this question (30% vs. 17% in 2015). While YouTube remains the most popular source of 'true and accurate information about fun things like hobbies and interests' (37%; unchanged since 2015), young people are now more likely to turn to Google for this (32% vs. 26% in 2015)' (Ibid: 86).

To investigate the existence of filter bubbles in young people's communication environments, I therefore focussed on search engines as a gateway to

knowledge. I used digital methods to capture search engine data and qualitative methods to place search engine use into the context of other platforms and modes of communication, including talk between peers, teachers, and family. I sought and received permission from the students to collect their web traffic and search histories (from pcs and laptops) during the two-hour window of my research while they connected to the college network and used its web server. Ideally, I would have collected these from the student's phones, tablets and laptops when they accessed the web in different contexts outside of college. However, once again, given the potential sensitivity of data generated by such intrusive research, ethical clearance would have been difficult to obtain.

Formulating topics to search

To operationalise this idea of search engines as a gateway to knowledge I wanted to give the students topics to research that interested them; topics that they would research 'normally' (outside contrived research conditions), and topics that would potentially mobilise differences in their culture and class. Therefore, at each site, I began with an hour-long semi-structured group interview, during which we discussed how they used the web for information and the topics that interested the students. Guided by previous research in this area by Robinson (2013), I hypothesised that their preconceived ideas of how search engines work would also influence the form and content of their searches. Therefore, I also used this opportunity to ask the students to reflect on how search engines retrieve and organise results.

Based on the topics that engaged the cohort in the group interviews, I gave the groups a series of contested or controversial questions to answer individually and without looking at any sources (off or online), before returning their answers to me on paper. Next, I asked the students to research these questions online. Although the

questions were informed by the group interview so they could reflect the students' concerns, I included some identical questions across both sites so I could directly compare search engine usage at each site, and the influence of the technologies of production on the search results. These questions included:

“Is there evidence for *any* conspiracy theories; about, for example, 9/11?”

“Does aspartame cause cancer?”

“Should we have more or fewer restrictions on immigration?”

“Did the last Labour government cause the crash?”

“Does smoking marijuana cause mental illness?”

I tested these questions at each research site to show that, depending on the key words or phrases the students used, they would see a range of disagreeing sources of information. I asked the students to support their answers with websites they considered credible. I told the students they could also, if they wished, discuss the questions within the group during the exercise. Each group had up to two hours to complete these tasks. During the exercise, the students at each site occupied the same room and I audio and video recorded all the proceedings, including all the intra-group discussions and interactions the students had while they were deliberating the questions. I also captured the students' search histories and search terms, and the search results they interacted with.

Capturing the searches

For the research exercises I was given a room of PCs at Chancery and Apple Macs at Calshot. I installed software called Squid (<http://www.squid-cache.org/>) on my laptop, which is designed for caching web traffic by proxy. Using this, I was able to capture in real time all the http traffic that was sent to each client machine the students used via the institution's web server and router. HTTP is an application layer protocol to

exchange or transfer hypertext on the Internet. When a user clicks on a link (or “URI”), the link contains an address on the internet for the data associated with that link, such as an HTML document representing a webpage. Web browsers send a request for the data, and the server (the computer on which the data is stored) sends the data to the client’s browser. The Squid software captured the addresses of all the client machines and data, and the data’s file extension. The data came as lines of URLs that detailed the client/server transactions. This included the date and time the data was sent.

Analysis

Did the technologies of production create filter bubbles?

We are told Google’s search engine reads a series of signals while it compiles its SERPs. At the “page level” Dean (2016) lists “key word density” and an array of “content quality signals” such as: “Latent Semantic Indexing Keywords in Content (LSI) (LSI keywords help search engines extract meaning from words with more than one meaning; for example Apple the computer company vs. the fruit)”; “recency of content updates”; “spelling and grammar”; “number of broken links”, “in bound link density”, and “Page Rank”. At the “site level” there is “server location”, “user reviews and site reputation”, “Page Rank of linking page”, “diversity of link types”, “link from authority sites”, “user interaction and repeat traffic”. Domain factors include the domain’s age and country code such as .uk We are then left with a series of “special algorithm rules” which take into account “user browsing history” and “user search history” to produce “search chain influence” or search results that are cached for later searches. For example, if a user searches for “reviews” then searches for “tablets”, we are told Google is more likely to show tablet review sites higher in the

SERPs. These various levels; server page level and so on, are not considered discrete. Rather all the factors can be used in combination. For example, Google is reported to use “geo-targeting” which means it gives preference to sites with a local server IP and country-specific domain name extension.

Yet, despite these signals, when the students used Google and the same key words, regardless of location (Calshot or Chancery), the search engine result pages (SERPs) were, remarkably similar. Although this study is on a relatively minor scale it supports the findings of Puschmann (2017) and Hannak et al. (2013) who tested 200 users on Google web search to find that, “on average, on average, 11.7% of results show differences due to personalisation” (ibid: 528), At both sites Wikipedia was prominent in both SERPs. Similarly, Cancer.org, an American website, and The Daily Mail’s website were conspicuous at both sites (see Tables 3).

Table 3:

The top four websites at each site

http://en.wikipedia.org/wiki/Aspartame_controversy
http://www.cancer.org
http://www.weightlossresources.co.uk
http://www.dailymail.co.uk/news/article-2290544/Aspartame-Cancer-premature-birth-fears-linked-fizzy-drink-sweetener.html

The data suggests the (material) technical effects of “bubbling” are relatively inconsequential, at least as far as search results are concerned. I was able to replicate the results on my laptop in various settings (at home for example) away from the research sites. If I typed in “Does aspartame cause cancer?” or the key words “aspartame” and “cancer” the results were very similar to the results produced at

Calshot College and the Chancery School, and being logged in or out of Google made little difference to the results. When investigating the aspartame question the only significant difference between sites is the website <http://aspartame.mercola.com>, which appeared in my own first SERP and in as the fifth ranked result on the Chancery School proxy logs. This is a commercial site that claims (counter to the scientific consensus) that aspartame is by “Far, the Most Dangerous Substance on the Market”. It is absent from the Calshot College logs. How can we account for this? It is difficult to implicate any of the 200 signals. However, Google announced in 2011 that its ranking algorithms would be influenced by social networking metrics (Google Official Blog, 2011). “Shares” within Google’s own social network, *Circles*, for example, would count within Google’s algorithms as significant votes or recommendations for the website’s content, thereby pushing the “shared” site up to the top of Google’s SERPs. The website is run by Dr Joseph Mercola, an osteopath whose empirically unsubstantiated claims about food as well as vaccines, HIV, and sunscreen have been discredited by scientific authorities such as the WHO and FDA. Nevertheless, Mercola’s popularity on Circles that was established in the three months that elapsed between my visit to Calshot and Chancery promoted him to the top of Google’s SERPs in the Chancery logs, where he remains at the time of writing. Mercola’s website is also missing from Omar’s log (Chancery) who used DuckDuckGo instead of Google to answer the aspartame question.

Yet on other topics there were significant differences between the content of the SERPs between sites, and between individuals at each site. For example, the Chancery students saw fewer anti-immigration websites in their SERPs than many of the Calshot students. What accounts for these differences?

Separating these technologies in forms that emerge from this case study including, students' ideas about search engines, their understanding of how search works, the epistemic conditions in which they had learned to use search engines, and idea about how society works emerging from a combination of their culture and class is difficult. For example, knowing that there are other search engines available that are designed to prevent filter bubbling, knowledge of Google's prioritising algorithms, informed scepticism about search results, and the knowledge and analytical skills to assess the accuracy of information are forms of technical and cultural capital (Halford & Savage, 2010) that (as this study will show) can be accumulated as a consequence of being a member of the professional class. However, I separate these ideas thematically to help structure the analysis. Previous research, by addressing filter bubbles as primarily a technical phenomenon, effectively operationalises technical determinism. The object of the current analysis, by contrast, is to demonstrate that the filter bubble is as much a social concern as a technical problem: if not more so.

Accounting for the differences - beyond material technological effects

As they often intersect and merge in the politics of search and interpretation, it is difficult to disaggregate the “technologies of sign systems” from the “technologies of power” (Foucault, 1988). I will therefore focus on the themes that emerge from thinking about these technologies as products of the social rather than the material: namely choice and understanding of search engine technology, the epistemic conditions of search, and culture and class.

Choice and understanding of search engine technology

By the study's end, Omar from the Chancery School had been accepted to Oxford to study computer science and philosophy. Omar's father is a computer scientist. Omar told me he had taught himself to code and that he participated in specialist online communities for programmers such as Stackoverflow. Omar said his parents had given him full autonomy online at home from a young age. The group interview data suggests that Omar had a reputation among his peer group as an expert user. For example:

Me: "Is there anyone in the school with a reputation for knowing more about IT than you, Omar?"

One of the young men mentions another name.

Omar: "He's quite close to me."

Tom: "Not anymore!"

There are laughs around the room that suggest Omar's reputation has been enhanced during this interview.

All these factors conferred on Omar forms of technical and cultural capital (Halford & Savage, 2010) that enabled him to critically engage with search engine technology: he was even aware of the discourse around filter bubbles. He revealed he avoided using Google and instead used DuckDuckGo because it promises in its mission statement to provide the antidote to so called "bubbling" by providing unadulterated SERPs. As he explained in the interview, after I asked him; "Has anyone noticed different results from Google if you are logged in? For example, if you read the Guardian a lot then it may come up on your first page of results?"

Omar replied: “Isn’t that called bubbling? When you’re only given information for things you are looking for? And so you end up seeing only stuff you are likely to agree with. If you’re doing research about something.”

Interviewer: “How do know about this?”

Omar: “Because I use DuckDuckGo as my main search engine and it has a campaign against bubbling on their website.”

Table 5 provides evidence from his history file of how Omar put this commitment to avoid bubbling into practice by using multiple search engines. Column 1 is the time the content arrived at his browser.

Table 4:

Omar trying to avoid bubbling by using multiple search engines

Time	Search engine and search terms
08:37:59	https://duckduckgo.com/debt+crisis
08:38:26	http://www.google.co.uk/search?site=&source=hp&q=Debt+Crisis
08:47:56	https://duckduckgo.com/?q=debt+crisis+UK+Labour
08:48:11	https://duckduckgo.com/?q=debt+crisis+UK
08:50:56	http://www.bing.com/search?q=debt+crisis+UK
08:58:07	https://duckduckgo.com/?q=labour%27s+financial+crisis+UK
08:58:38	https://duckduckgo.com/?q=financial+crisis+UK+labour
08:58:51	https://duckduckgo.com/?q=financial+crisis+labour
08:59:06	https://duckduckgo.com/?q=debt+crisis+labour
08:59:58	https://duckduckgo.com/?q=UK+debt+crisis+labour

In groups Omar was anxious to reinforce his identity as a critical technical expert. This led to him distributing his technical and cultural capital throughout his peer group. His statements about technology became ‘facts’ in the sense they were not, at least publicly, contested by the group and his statements influenced the practice of his peers. For example, the proxy logs show other students using DuckDuckGo as the research session progressed. When the students used DuckDuckGo for the same topics they saw different content in their SERPs to students who used Google. The proxy logs show everyone at Calshot using Google to answer their questions, with the college’s browser defaulting to the college’s homepage. To perform a search the students entered google.com or google.co.uk into their browser’s address bar. In interviews the Calshot students described searching the web as “Googling”.

Comparing SERPs between search engines would have been a useful educational exercise for the Calshot students, perhaps increasing their awareness of search engine retrieval and how algorithms prioritise returns. However, they used only Google. In fact, neither group understood how search engines prioritise results. When I initially questioned the Chancery students about how Google works, none of young men apart from Omar was prepared to offer an answer. He referred to PageRank and this then became accepted knowledge within group. For example, in the research exercise, when I asked the students “How does Google prioritise its search results?” it was evident from the proxy and history logs that almost all the students looked for “PageRank”. The Chancery students therefore learned about search engines during the exercise. There was no equivalent of Omar at Calshot. None of the students at Calshot knew about DuckDuckGo, nor could they offer an explanation of how Google retrieves or prioritises its search results. As Sarah said:

“I think the order is random: I don’t know how Google works”.

“No one’s ever taught us about anything else”, added Lizzy.

Knowledge and understanding of search engine technology was a form of technical and cultural capital that the Chancery students could access but that was unavailable to the Calshot students.

The epistemic conditions of searching for information

Given the nature of the questions they researched, I expected that the students’ search terms and results would vary. However, at Calshot, as evidenced in the proxy logs, for many of the questions all the students entered the whole question into Google. I had (thinking back on it) facilitated this technique by providing the students with a digital worksheet from which they could copy and paste the questions from Word into Google. During the group and individual interviews, however, the students confirmed that this was something they do normally, and that had I given them a paper sheet they would have just typed the whole question in anyway. Liam said if he does use key words it is for “quick searches”; about football for instance, but if someone asked him to answer a question he would “type in the whole question into Google”. This is indicative of a distinction that emerged at Calshot between using key words in leisure-orientated searches and typing out specific questions for educational purposes (or what the students called “college work”). The study’s individual interviews took place during a class within which the students had been asked by their tutor: “How do you start up your own business?” This practice, of typing out the question, was confirmed by Alfie (as well as others) *in situ*:

“If I’m looking for a song I’ll type a few words from the song but if I’m asked how to set up a small business; better to ask Google a question rather than type in ‘small business’”.

“I learnt that myself”, he added.

Sarah also corroborated what her peers told me:

“I type in keywords, but if it’s work I find it a lot more helpful if I type in the question”. She said she “Couldn’t remember doing anything different”.

Calshot, as a post-16 higher education college, recruited students from schools in the area that had no sixth form (post 16 provision). The Calshot students mainly came from one of three state schools in the area yet typing or copying in the whole question was common practice among the cohort. This suggests they were treating Google as an authority that could understand and answer questions the students would normally ask teachers - something that had emerged from the classed and gendered habits, dispositions, and expectations that were normalised in all the Calshot students’ educational experiences (Davies, 2017).

The practice of typing in the whole question was somewhat frowned upon by the Chancery School students as a less skillful way of interrogating Google. The students there associated more sophisticated search with different combinations of key words. This suggests that they regarded proficiency with a search engine as another form of technical and cultural capital. Contrary to self-reports, however, some of the Chancery students also expected Google to answer ‘whole’ questions. This exposes the limitations of using self-reporting to assess a user’s skills (as discussed, for example, in Litt 2013). When the Chancery students were asked; “How does how does Google prioritise its search results?” even Omar, despite his reputation and self-confidence in his expertise typed in the whole question: “how+does+google+prioritize+its+search+results”. However, Google is increasingly applying NLP (natural language processing) to its search (Google, 2017) so that key word searches and their associated strategies are becoming obsolete; particularly as talking to an Internet-connected device is becoming much more conventional.

When the students researched some of the other questions, especially about politics, an important distinction emerged between the Calshot students, who needed Google to give them information about topics they were researching for the first time (such as the 2008 economic crash), and the students at Chancery who could draw on other forms of cultural capital that were circulating through their lives, including studying politics in school, reading broadsheet newspapers and political magazines, and talking to family and friends. This meant they didn't have to rely on search engines, or that they knew exactly where to find answers to the questions. These, for example, were Ryan's (Chancery) search terms.

Table 5:

Ryan's search log that shows he performed only five discrete searches in an hour.

09:01:17.804	http://www.google.co.uk/search?q=how+much+money+is+spent+on+welfare+in+
09:32:51.668	http://www.google.co.uk/search?q=global+warming+hoax
09:33:08.547	http://www.google.co.uk/search?q=immigration+uk&oq=immigration+uk
09:49:55.121	http://www.google.co.uk/search?q=aspartame
10:02:19.533	http://www.google.co.uk/search?q=how+does+google+prioritise+results

Unlike the Calshot students, who often copied and pasted their answers from websites that they considered to have answered the question, Ryan didn't have to use content from websites to answer the political questions. He called on his cultural capital to argue that the last Labour government didn't cause the crash. This an extract from an answer of over 500 words:

“Had Labour not recapitalised the banks, there would have been a catastrophic global depression, causing much higher levels of debt and, indeed, destitution as

people would have lost all their savings. Therefore, the last Labour government is not to blame; rather the deregulation of the financial sector and the failure of the neoliberal economic model to provide adequate employment. Despite this unequivocal fact, right wing politicians have sought to blame “dependent” people on welfare for our debt woes, thus moving society even further in the direction of capital.”

The data suggests that relative dependency on search engine to give definitive answers is another important signifier of (lack of) cultural capital.

Culture and class

The students’ response to immigration, in particular the discourse they used, speaks of each cohort’s relative position in the political economy and how they explain it. The individuals in the Calshot cohort assembled a mosaic news stories, social media memes, and anecdotes and inventions shared by their families and peers to produce a story that positioned themselves as the indigenous victims of a political conspiracy that privileges immigrants at their expense. As William said, “We can’t afford houses yet we pay for *their* big houses in Westminster and central London”. Liam had read and heard similar stories: “There’s this like single woman from wherever, I dunno where’s she’s from, she’s from Afghanistan or wherever like that, she got eight children she lives in a four million pound mansion in London”. They had a strict normative definition of an immigrant as someone who was unskilled, lived on benefits or, at best, semi-skilled. Even the Calshot students who were relatively sanguine about immigration, such as Jess, didn’t regard immigrants as potential professionals. She said they do the ‘dirty jobs’ that British people refuse to do. The majority of the Calshot students looked on immigrants as competitors for jobs and

housing who received resources they said immigrants didn't 'deserve'. This was reflected in their search terms. These are George's, for example:

Table 6:

George's (Calshot) search terms to answer the question; "Should we have more or fewer restrictions on immigration?"

'how much do immigrants cost the NHS?'
'where do most immigrants come from'
'Immigrant mansions'
'what is the cost of immigration'
'how much benefits do immigrants get?'

Many of the Chancery students' parents were first or second generation immigrants who worked in highly skilled jobs. This group was typical of the whole school, which combined white middle and upper class British and children of professional immigrants. Highly skilled immigrants were therefore Chancery students' peer group, their in group, and their imagined community. The Chancery students reinforced their identity and value beneficiaries by advocating a globalised open market within which immigrants could monetise their skills as a public good that benefits the economy. For example, Ryan argued that 'in the EU it greatly aids business'. He added: 'An immigration cap would slam the door in the faces of skilled workers who help our economy and further scientific progress.' This position was also reflected in the search terms the students used at each site – compare Table 6 George's search at Calshot to Table 7 Zac's search at Chancery.

Table 7:

Zac's (Chancery) search terms to answer the question; "Should we have more or fewer restrictions on immigration?"

"reasons for immigration"
"reasons supporting immigration",
"immigration economic benefits"
"pro-immigration arguments"

This is only a very small case study. Google owns billions upon billions of such individual searches that could provide much more generalisable data but Google keeps its data in house. Google Trends allows us to see such data at the level of a region or city but nothing smaller. Nevertheless, from this small study can see the socio-technical recursion I referred to in this article's title. A recursive algorithm calls upon its own output values (or outputs from its sub routines) to continue without interruption by (re)integrating its own outputs into its routines and adapt to its own functioning (Jordan, 2015). Within a recursive programme; "something the user has, in the information they enter, is taken by the recursive programme and served-up to the controller of that programme. This transfer benefits the ongoing work of the programme, thereby benefitting the controller of the programme" (ibdi:40). Socio-technical recursion - users coming back for more because they like the suggestions offered to them which have been inferred from their previous searches - benefits the owners of the platform. The concept of the socio-technical system originates from STS (Trist, Emery, & Churchman, 1960) to describe the intricacies of interactions between socially situated actors and the organisational structures they inhabit. Actor Network Theory (ANT) subsequently denied any distinction between the social and the technical in its use of the term. Because, however, ANT also refutes 'any a priori

assumptions about the existence of social categories or aggregates' (Halford & Savage, 2010: 938) it has little to contribute to the analysis in this article. To locate the origins of the assumptions that feed this recursion we have to look at an individual's whole information ecosystem including every source of information; from friends and family to news media and social network feeds. However, to break this recursion users require the necessary forms of cultural and technical capital to challenge the message *and* the medium. Hall (1980) offers us a theoretical foundation for this challenge.

Conclusion

In the days of non-interactive analogue television, filter bubbles around visual content were a consequence of having only a few channels to choose from. At the time, Hall (1980) invited us to consider the complex relationship between culture, class, and television by suggesting that its consumers pre-encode its content so that the "audience is both the 'source' and the 'receiver' of the television message" (ibid: 92). Circulation and reception of content were "moments of the production process [that were] reincorporated into the structured feedbacks of the production process" (ibid: 93). But "production and reception" were not thought of as identical: they were "differentiated moments within the totality formed by the social relations of the communicative process as a whole". Once produced, the message was "meaningfully decoded" (ibid: 93) within the "politics of signification" (ibid: 103). Conceptually, this is similar to the Foucauldian definition of technology I have used here. Hall implies that there is little meaningful distinction between immaterial and material technologies and that the medium and the message are defined by the receiver's relationships to power.

When Hall was writing there were few ways of accessing the thoughts of television audiences: for the producers of television their audience was primarily composed of imagined communities that were inferred from viewer demographics, focus groups, and letters of complaint, rather than individuals who could influence the content of the broadcast text in real time. New media platforms have closed the gap between the encoder-producer and the decoder-receiver so that the audience co-produces (in real time) the content they generate. Hall's analysis is therefore more relevant than ever: pre-encoding takes place in the production and the interpretation of the text.

The data here suggests that socio-technical recursion is the result of the culturally mediated and classed politics of signification, which Google and other media platforms such as Facebook and Twitter participate in without acknowledging their contribution (Bell & Owen, 2017). These platforms share a broadly similar libertarian worldview, which, in practice, means all ideas deserve an equal podium. As Mark Zuckerberg wrote: "That's what running a platform for all ideas looks like" (Zuckerberg, 2017). This libertarian worldview is one of Silicon Valley's technologies of power. It has led to, for example, Google creating a controversy for allowing Holocaust denial websites to rank highly in its SERPs (Griffin, 2016). These platforms are not neutral in the ways that they owners claim they are. They have "evolved beyond their role as distribution channels, and now control what audiences see and who gets paid for their attention, and even what format and type of journalism flourishes" (Bell & Owen, 2017: 9): they don't just present 'facts': they present messages about facts. Google's preferred hegemonic reading for its users is that it provides the best results available that best reflect your search terms, or that are the

best-calculated answer to your question. This gives Google's SERPs what Hall (1980) called a "stamp of legitimacy" (ibid: 102).

The young people in the two student cohorts discussed here pre-encoded their searches with the search results they wanted to see. Some, if they failed to find something that validated their views, would try different search terms until they agreed with (this is shown in relationship between the websites the students choose to represent the best answer and the sequence of their search terms). It can be a mistake therefore to assume there is a distinctive event when an individual is convinced by a source to change her mind. Pew found that only 56 per cent of online news consumers who had clicked on a link could recall the news source and only 20% of online readers said they had changed their mind about an issue after reading about it on social media (Duggan, Smith, & Page, 2016). Hall said the event must become a 'story' before it can become a communicative event (ibid: 92). Search engines can be a way of authenticating that story: in this study, anti-immigration searches produced anti-immigration websites. This produces what Hall calls a "fundamental alignment and reciprocity – an achieved equivalence – between the encoding and decoding sides of an exchange of meanings" (ibid: 95). As work by Baker & Potts (2013) and more extensively Noble (2018) shows, this can produce a self-reinforcing "prejudice in, prejudice out" model of media engagement, which is the fear of the filter bubble realised.

These media and search platforms can be gamed by actors who are able to get their communicative event prioritised by ranking algorithms so that 'the stamp of legitimacy' becomes a method for delivering political propaganda (Albright, 2017). This "strategic distortion of attention can hasten the spread of misinformation and the establishment of 'alternative facts'" (ibid: 88). When users see these alternative facts

in their homophilous networks they are exposed to “the majority illusion” (Lerman, Yan, & Wu, 2015); the idea that everyone in their community (actual and/or imagined) shares their views.

The recursion metaphor suggests I have substituted technological determinism for socio-technical determinism. Yet, many of the young people in this study were not equipped with cultural and technical capital to decode Google’s “stamp of legitimacy”, and some of them therefore had their prejudices validated by a search engine’s results. Google is beginning to engage with this problem; its parent company’s chairman Eric Schmidt has said Google engineers are working on ‘de-ranking’ Russia Today and Sputnik (BBC, 2017). But because they are socio-technical accomplishments there are many layers of structurally shaped human experience that go into their creation. Technology could be part of the solution but there is no technical fix for socio-technical recursion.

To help escape this form of socio-technical recursion, we have to challenge the preferred reading that “platform capitalism” (Srnicek, 2016) has reserved for its platforms as epistemically neutral information delivery systems: they are material and political technologies of power. This begins by foregrounding and investing in academic disciplines that are developing interdisciplinary, critical, and analytical frameworks that account for these effects: then distributing their findings for greater public awareness. Sociology is well equipped to address three out of four Foucault’s technologies (Foucault, 1988). Its contemporary challenge is to identify, understand and critique the recursive computational thinking that is encoded with Silicon Valley’s worldview. It shows us what to watch on YouTube or finishes our search queries on Google. These platforms expose us to ‘all ideas’ then they monetises our attention to consolidate their market share and power. This critical approach to

technology then needs to become a form of technical and cultural capital that is available to all of us, not just a fortunate few who study it for a living.

- Albright, J. (2017). Welcome to the Era of Fake News. *Media and Communication*, 5(2), 87. <http://doi.org/10.17645/mac.v5i2.977>
- Baker, P., & Potts, A. (2013). “Why do white people have thin lips?” Google and the perpetuation of stereotypes via auto-complete search forms. *Critical Discourse Studies*, 10(2), 187–204. <http://doi.org/10.1080/17405904.2012.744320>
- Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, 348(6239), 1130–1132. <http://doi.org/10.1126/science.aaa1160>
- BBC. (2017). Google to “derank” Russia Today and Sputnik. Retrieved November 21, 2017, from <http://www.bbc.co.uk/news/technology-42065644>
- Beer, D. (2009). Power through the algorithm? Participatory web cultures and the technological unconscious. *New Media and Society*, 11(6), 985–1002. <http://doi.org/10.1177/1461444809336551>
- Behrent, M. C. (2013). Foucault and Technology. *History and Technology*, 29(1), 54–104. <http://doi.org/10.1080/07341512.2013.780351>
- Bell, E., & Owen, T. (2017). The Platform Press: How Silicon Valley Reengineered Journalism, 94. <http://doi.org/10.7916/D8R216ZZ>
- Chadwick, A. (2013). *The Hybrid Media System: Politics and Power*. Oxford and New York: Oxford University Press.
- Davies, H. C. (2017). Learning to Google: Understanding classed and gendered practices when young people use the Internet for research. *New Media & Society*, 1461444817732326. <http://doi.org/10.1177/1461444817732326>
- Duggan, M., Smith, A., & Page, D. (2016). The Political Environment on Social Media. Pew Research Center, 38. Retrieved from

<http://www.pewinternet.org/2016/10/25/the-political-environment-on-social-media/>

Dutton, W. H., Reisdorf, B. C., Dubois, E., & Blank, G. (2017). Social Shaping of the Politics of Internet Search and Networking : Moving Beyond Filter Bubbles , Echo Chambers , and Fake News Search and Politics :, (September).

<http://doi.org/10.2139/ssrn.2944191>

Erola, J., Jalonon, S., & Lehti, H. (2016). Parental education, class and income over early life course and children's achievement. *Research in Social Stratification and Mobility*, 44, 33–43. <http://doi.org/10.1016/j.rssm.2016.01.003>

Flaxman, S. R., Geol, S., & Rao, J. M. (2016). Filter Bubbles , Echo Chambers , and Online News Consumption. *Public Opinion Quarterly*, 80(Special Issue), 298–320.

Foucault, M. (1988). Michel Foucault, Technologies of the Self. In L. H. Martin (Ed.), *Technologies of the Self: A Seminar with Michel Foucault* (pp. 16–49). University Massachusetts Press.

Google. (2017). Search more intuitively using natural language processing in Google Cloud Search. Retrieved November 21, 2017, from <https://www.blog.google/products/g-suite/search-more-intuitively-using-natural-language-processing-google-cloud-search/>

Griffin, A. (2016). Google results claim The Holocaust didn't happen, and the company won't change it. Retrieved November 15, 2017, from <http://www.independent.co.uk/life-style/gadgets-and-tech/news/google-holocaust-did-happen-results-search-shoah-anti-semitism-stormfront-a7477041.html>

Halford, S., & Savage, M. (2010). Reconceptualizing digital social inequality,

- (918675184). <http://doi.org/10.1080/1369118X.2010.499956>
- Hall, S. (1980). Encoding/decoding. In S. Hall, D. Hobson, A. Lowe, & P. Willis (Eds.), *Culture, media, language-Working papers in cultural studies, 1972–79*. (pp. 117–127). London: Routledge. <http://doi.org/10.1007/BF00986815>
- Hannak, A., Sapiezynski, P., Arash Molavi Kakhki, Balachander, K., Lazer, D., Mislove, A., & Wilson, C. (2013). Measuring Personalization of Web Search. In *Proceedings of the 22nd International Conference on World Wide Web* (pp. 527–38). Retrieved from <http://dl.acm.org/citation.cfm?id=2488435>
- Jordan, T. (2015). *Information Politics*. London: Pluto Press.
- Lerman, K., Yan, X., & Wu, X.-Z. (2015). The Majority Illusion in Social Networks, 90292. <http://doi.org/10.1371/journal.pone.0147617>
- Litt, E. (2013). Measuring users' internet skills: A review of past assessments and a look toward the future. *New Media & Society*, 15(4), 612–630. <http://doi.org/10.1177/1461444813475424>
- McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, 27(1), 415–444. <http://doi.org/10.1146/annurev.soc.27.1.415>
- Mitchell, B. Y. A., Gottfried, J., Shearer, E., & Lu, K. (2017). How American Encounter, Recall and Act Upon Digital News. *PewResearchCenter*, 1–41.
- Negroponte, N. (1996). *Being Digital*. Coronet Books.
- Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press.
- Ofcom. (2013). Children and Parents: Media Use and Attitudes Report, (October), 1–53. Retrieved from <http://stakeholders.ofcom.org.uk/binaries/research/media-literacy/october-2013/research07Oct2013.pdf>

- Pariser, E. (2011). *The Filter Bubble: What The Internet Is Hiding From You*. Penguin.
- Puschmann, C. (2017). How significant is algorithmic personalization in searches for political parties and candidates?
- Research Centre, T. P. (2016). *The Modern News Consumer*. Retrieved from http://www.journalism.org/files/2016/07/PJ_2016.07.07_Modern-News-Consumer_FINAL.pdf
- Robinson, L. (2013). Freeways, detours, and dead ends: Search journeys among disadvantaged youth. *New Media & Society*, 16(2), 234–251. <http://doi.org/10.1177/1461444813481197>
- Srnicek, N. (2016). *Platform Capitalism* (Kindle). Polity Press.
- Steinberg, T. (2016). No Title. Retrieved August 1, 2016, from <https://www.facebook.com/tom.steinberg.503/posts/10157028566365237?pnref=story>
- Sunstein, C. R. (2006). *Infotopia: How Many Minds Produce Knowledge*. Oxford University Press.
- Trist, F. E., Emery, E. L., & Churchman, C. W. (1960). Socio-technical systems, Management Science Models and Techniques. *Management Science Models and Techniques*, 2, 83–97.
- Zuckerberg, M. (2017). Facebook's response to President Trump's claims. Retrieved November 15, 2017, from https://m.facebook.com/story.php?story_fbid=10104067130714241&id=4
- Zuiderveen Borgesius, F. J., Trilling, D., Möller, J., Bodó, B., de Vreese, C. H., & Helberger, N. (2016). Should we worry about filter bubbles? *Internet Policy Review*, 5(1), 1–16. <http://doi.org/10.14763/2016.1.401>

